Department of Agriculture
Soil Conservation Service
Central Technical Unit

ELEMENTS FOR DETECTING OUTLIERS IN SAMPLES FROM A KNOWN CUMULATIVE PROBABILITY DISTRIBUTION

by
William H. Sammons

September 1971
Hyattsville, Maryland

ELEMENTS FOR DETECTING OUTLIERS IN SAMPLES FROM A KNOWN CUMULATIVE PROBABILITY DISTRIBUTION

William H. Sammons[1/]

INTRODUCTION

Hydrological data are collected usually for purposes of interpretation and if proper use is to be made of the information thus obtained, then some decision should be reached or some action taken as a result of analyzing the data.

In many cases a critical examination of the data collected is necessary in order to insure that the results of sampling are representative of the thing or process under examination. Quite frequently our observations do not appear to be consistent with one another; i.e., the data may seem to display non-homogeneities and the group of observations as a whole may not appear to represent a random sample from say, a single population or universe. In particular, one or more of the observations may have the appearance of being "outliers." It must be determined whether such observations should be retained in the sample for interpreting results or whether they should be regarded as being inconisstent with the remaining observations and censored. Actually, the rejection of "outlying" observations may be just as much a practical (or common sense) problem as a statistical one and sometimes the practical or experimental viewpoint may naturally outweigh any statistical contributions. The final analysis would seem to reduce to a question of common sense. Certainly the judgment of an experienced hydrologist should be allowed considerable influence in reaching a decision. The outcome is to be aided by the application of one or more tests based on the theory of probability, etc. See References 1, 2, and 3.

Most of the statistical tests that have been recommended for the detection of a single (or more than one) outlier involve the difference between the largest (or smallest) observation of a normal or standard normal distribution and some measure of the location of the remaining members of the sample. A skewed distribution will require first, a transformation (graphical and/or analytical). In many situations, there is no a priori basis for anticipating which extreme (or extremes) will be under suspicion with the result that this decision is based entirely on sample evidence. An example could be the sub-set of a sample which indicates one (or more) outliers although the whole sample does not. The test statistic in such cases actually employs

[1/]Hydraulic Engineer, Central Technical Unit, Hydrology Branch
Soil Conservation Service, Hyattsville, Maryland.

the "more deviant" extreme, and this reordering is seldom taken into account. Hence, it would appear that statistical tests of significance for judging or testing "outliers" comes into importance either in supporting doubtful practical viewpoints or in providing a basis for action in the absence of sufficient experimental knowledge of underlying causes in an investigation. In this paper the sample statistics are assumed to be efficient and defined by the observations (after a proper transformation of the original observations to the normal case). Only the largest (or smallest) observation in the sample will be considered since the methods for two or more large (or small) observations are not well defined at this time. The process will therefore be of a repetitive nature if more than one large (or small) outlier is present.

Briefly, the process is as follows: First, each sample is to be ordered according to magnitude and plotted on log-normal probability paper, (Codex Book Co., Inc. No. 31,376 or equivalent). (A curve is fitted to this plotted data). Second, the standard normal deviate (K_n value) will be read from the curve on the linear scale for the largest and smallest values of the sample. Last, the standard normal deviate (K_n value) obtained will be compared with the "critical values" tabulated in Tables 1 and 2 for selected percentage rejection levels. Then the process may or may not be repeated to determine if additional outliers are present --- only considerable experience can be substituted for this repetitious process.

Methods of Analysis

In Reference 4 the following quote should be noted:

Maximum likelihood estimation (MLE) (17), best linear unbiased estimation (BLUE) (33,34,35), best linear invariant estimation (BLIE) (35,36) and so on, with and without single and double censoring, should be employed. The usual method of moments estimation (MME) is too inefficient. No outliers are recognized in a population. Outliers may be present in a mixture of populations, or they may be present in a sample from a population or a mixture of populations. For no other reason, except lack of knowledge in the estimation process, outliers must be censored in order to estimate reasonable statistics for the sample. Graphical work cannot be used unless one is willing to assume a probability distribution function (pdf) and cumulative distribution function (cdf), or both, where outliers are to be considered. The first step in any analysis is to assume a pdf or cdf in a synergistic way - based on experience. Shotgun methods may be used in initial studies of an area. Seldom should one expect a single sample or station to represent an area or region. The writers have used Grubbs' (25) and Wilks' (44) method of identifying outliers when censoring the data. Two computer programs were used both of a multipurpose nature, giving MME and MLE solutions. The BLIE solutions were manually computed since the computer programs have not been completed.

Normal Distribution

This is the basic distribution and all other distributions maybe referred back to the normal. Compute the arithmetic mean ($\bar{X}$), standard deviation (S_x) and/or Coefficient of Variation (C_v) from the sample data. The equation is:

$$X = \bar{X} \pm K_n S_x \quad (1)$$

or

$$(X/\bar{X}) = (1 \pm K_n C_v) \quad (2)$$

where

$$K_n = (X - \bar{X})/S_x \quad (3)$$

K_n is the normal deviate and/or reduced variate. Tables 1 and 2 - "Critical Values, Probabilities and Return Periods for Two-tail Test for Detecting Outliers at 10 (and 5) Percent Level" contains the critical values of K_n.

Figure 1 has 4 examples based on a sample of 100. The curves are designated as (a), (b), (c) and (d). Curve (d) is the bi normal. Numbered points 1, 2, 3, 4, 5, 6, and 7 plotted at -2.50759 = K_n and 8, 9 and 10 plotted at +2.50759 = K_n are used to illustrate the detection process i.e., in (c) point 1 is not an outlier but point 2 is an outlier on the low end of (c). Point 8 is a high outlier but point 9 is not an outlier, etc. The following tabulation gives the mean and standard deviation for curves (a), (b) and (c):

	$\bar{X}$	S_x
(a)	5000.0	1618.001334
(b)	7000.0	970.8008004
(c)	9000.0	323.6002668

(d) Will be discussed under Broken - Line Frequency Distribution section.

Substitution of the mean and standard deviation into Equation 3 with the maximum or minimum observed data will give a K-value which is to be compared with the tabular values in tables 1 and 2.

Rule For Maximum and/or Minimum value:

> Computed K-Value should be less than the tabulated value of K for a given sample size.

If the computed value is equal to or greater than the tabulated value of K for a given sample size, then the observation is an outlier for this sample from normal distribution.

Log Normal Distribution (Zero Skewness)

In this case the log X = Y, where Y is a normal distribution. X has a skewed distribution. Compute the log mean (lg G), and the log standard deviation (lg S). The equation is:

$$Y = \log X = \log G \pm K_n \log S \qquad (4)$$

or the equivalent power form is

$$X = G\ S^{\pm K_n} \qquad (4^1)$$

where G is the geometric mean
S is the $\log^{-1}$ S (antilog S)

Equation 3 is now written as

$$\pm K = (\log X - \log G)/\log S \qquad (5)$$

and the same Rule applied for this example in Figure 2 as was applied to Figure 1 examples, i.e., 6 curves are given - (a), (b), (c), (d), (e), and (f). (f) is a bi log normal and will be discussed later. Curve (e) lower end, point 1 is not an outlier but point 2 is a low outlier - point 7 is a high outlier but point 8 is not when considering curve (e). Similar examples may be shown for the other 5 curves on this figure. The log mean and log standard deviation for these 5 curves, (a) through (e) are:

	log G	log S
(a)	1.500000	0.16180013340
(b)	2.000000	0.3236002668
(c)	2.500000	0.4854004002
(d)	3.000000	0.3236002668
(e)	3.500000	0.16180013340

Curve (f) will be discussed under Broken-Line Frequency Distribution section.

2 or 3 Statistic Gamma Distribution (Pearson Type III)

The Gamma distribution is based on 2 or 3 statistics as follows:

2-Statistics

$$X = U\ \beta\sqrt{\gamma} \qquad (6)$$

X is the original data
U is a function of γ and the probability level
β is the scale and is equal to $(\bar{X}/\gamma)$
γ is the shape and is a function of $\bar{X}$ and G

$\bar{X}$ is the arithmetic mean
G is the geometric mean
$\beta\sqrt{\gamma}$ is the standard deviation in original units

3-Statistics

$$X = U\,\beta\sqrt{\gamma} + C \tag{7}$$

C is the location
Equations 6 and 7 are identical if the original data goes to zero (or can go to zero). Ordinarily, C is some value equal to or less than the minimum observation in the sample data.

The above should not be confused with the situation where the period of record has a number of zero values.

In this case the probability of zero and non zero should be computed separately - their sum is equal to one for any distribution.

Below is tabulated the statistics of the 10 curves (a) through (j) on Figure 3. Beta was held constant at 48.924.

	γ	C
(a)	1	0
(b)	1	20.0
(c)	3	0
(d)	3	60.0
(e)	9	0
(f)	20	0
(g)	40	0
(h)	165	0
(i)	2	0
(j)	2	50.0

Figure 3 has been worked up for a sample size of 50 only to add variety and to show the effect of change of sample size. Curve (h) is a straight line for all practical uses and the log normal section could be applied to this one curve. Curve (a) demonstrates the steepest and most variable situations where as curve (b) with C=20.0 demonstrates the "S" shaped curve of the 3-statistic type. The Rule for detecting the presence of outliers is the same here since we use the K_n Value scale on the top of the log normal probability paper to get a graphical solution. For a mathematical solution the U-value is converted to the K-value which is then compared with the tabulated critical values of K at a given sample size. K_p values for the log-Pearson Type III are related to U-values and the Chi-square distribution for a given skewness, K_p values can be converted to K_n values which are compared with the critical values at a given sample size.

"Broken-Line" Frequency Distribution

W. D. Potter and many other "rugged individuals" have advocated two or more straight lines to fit the data. The result could be called bi normal, bi log normal, bi extreme value, etc.

Figure 1, curve (d) is an example of a bi normal where:

$X = (\bar{X}_1 \pm K_n S_{x_1}) + (\bar{X}_2 \pm K_n S_{x_2})$

for K_n $-\infty$ to 0.0 and 0.0 to $+\infty$

In terms of probability, we could write

$P(X) = \lambda P_1(X) + (1-\lambda)P_2(X)$

for bi normal, bi log normal and bi extreme value distributions

Figure 2, curve (f) is an example of a bi log normal where:

$\log X = (\log G_1 \pm K_n \log S_1) + (\log G_2 \pm K_n \log S_2)$

for K_n $-\infty$ to -0.52440051 and -0.52440051 to $+\infty$

Below is a tabulation of the statistics computed from a chemical analysis investigation using maximum likelihood estimation for the statistics (MLE.)

Examples of Statistics (MLE)

Set	N	Law	λ	m_1	σ_1	m_2	σ_2
Sr	143	BLN	0.468	462	0.497	61.2	1.04
		BN	0.655	430	262	47.7	317
K	80	BLN	0.107	2.63	0.080	3.89	0.103
		BN	0.0938	2.60	0.177	3.90	0.412
Na	134	BLN	0.172	0.137	0.632	2.22	0.290
		BN	0.149	0.128	0.060	2.27	0.752
Mg	120	BLN	0.0895	0.0342	0.448	0.255	0.639
		BN	0.825	0.225	0.132	0.591	0.335
K	140	BLN	0.148	2.50	0.372	4.44	0.169
		BN	0.081	2.08	0.463	4.42	0.871
Cl	127	BLN	0.376	151	0.963	492	0.352
		BN	0.162	106	531	472	231
SiO_2	440	BLN	0.595	66.1	0.0730	73.4	0.290
		BN	0.636	66.6	481	73.6	2.02

BLN is bi log normal law
BN is bi normal law
N is sample size
m_1 is a mean or geometric mean (original Units)
m_2 is a mean or geometric mean (original Units)
σ_1 and σ_2 is the standard deviation in original and natural log units

Conclusions

Any data analysis worth the time and effort is just an estimate of the population statistics, no more, no less. A single station may not define a region and may or may not be representative of the period of record in a regional study. There may be outlier stations in a region, just as there may be outliers in a sample set of data - censoring may be required in both cases. S. S. Wilks discusses outliers in k-dimensions - few investigators have absorbed the meaning of this final contribution from Dr. Wilks (died shortly after article was published). We all have heard the old saying that you can't learn to swim on dry land - well, you can't learn about outliers until you have made a few comparisons with and without same. There is a certain amount of art needed to help deal with the theory - rules used blindly can get you in to a lot of trouble - hydrology is involved and decisions must be made - not covered by rules - there is no substitute for experience.

References

1. Grubbs, Frank E., "Sample Criteria For Testing Outlying Observations," Annals of Mathematical Statistics. Vol. 21, No. 1, March 1950, pp 27-58.

2. Wilks, S. S., "Multivariate Statistical Outliers", Sankhya, Series A, 25(4) 1964, pp 407-426.

3. Sammons, William H., "Sample Criteria For Identification of Outlying Observations", March 1962 (In Service Publication).

4. Brown, Thomas L. and William H. Sammons, Discussion: "Flood Peaks From Small Southwest Range Watershed" by Kenneth G. Renard, John C. Drissel and Herbert B. Osborn. ASCE Proc. Paper No. 7161, March 1970, ASCE Proc. Paper No. 7783, January 1971 pp 182-193.

TABLE 1 --Critical Values, Probabilities and Return Periods for Two-tail Test for Detecting Outliers at 10 Percent Level

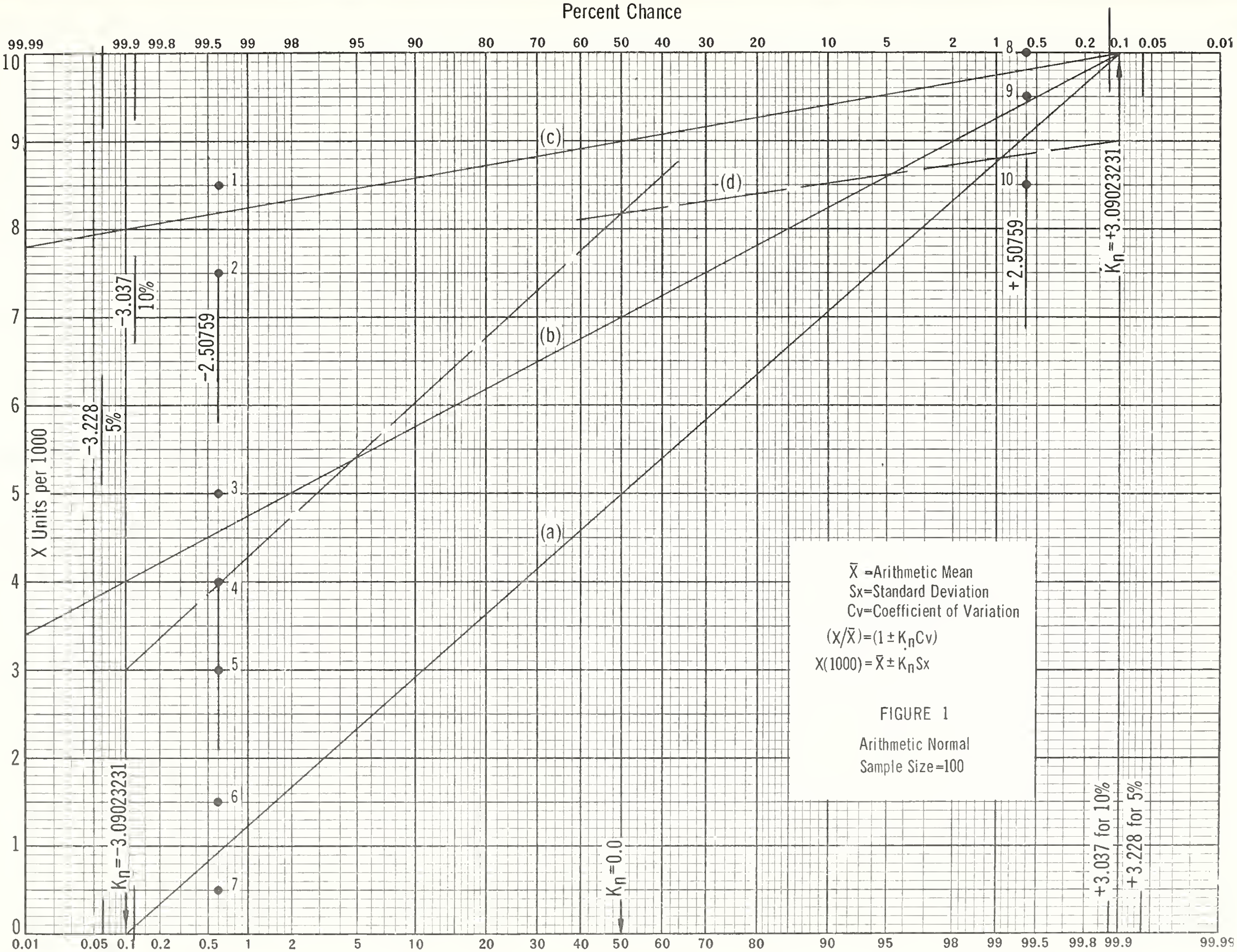
N	K10	P10	Q10	TP10	TQ10
4	1.645	0.9500151	0.0499849	20.00603	1.05261
5	1.791	0.9633534	0.0366466	27.28764	1.03804
6	1.894	0.9708875	0.0291125	34.34955	1.02998
7	1.974	0.9758092	0.0241908	41.33804	1.02479
8	2.041	0.9793746	0.0206254	48.48401	1.02106
9	2.097	0.9820033	0.0179967	55.56564	1.01833
10	2.146	0.9840636	0.0159364	62.74928	1.01619
11	2.190	0.9857380	0.0142620	70.11629	1.01447
12	2.229	0.9870931	0.0129069	77.47787	1.01307
13	2.264	0.9882130	0.0117870	84.83925	1.01193
14	2.297	0.9891906	0.0108094	92.51239	1.01093
15	2.326	0.9899908	0.0100092	99.90779	1.01011
16	2.354	0.9907137	0.0092863	107.68568	1.00937
17	2.380	0.9913437	0.0086563	115.52249	1.00873
18	2.404	0.9918916	0.0081084	123.32922	1.00817
19	2.426	0.9923669	0.0076331	131.00853	1.00769
20	2.447	0.9927975	0.0072025	138.84056	1.00725
21	2.467	0.9931875	0.0068125	146.78871	1.00686
22	2.486	0.9935406	0.0064594	154.81277	1.00650
23	2.504	0.9938601	0.0061399	162.86977	1.00618
24	2.520	0.9941323	0.0058677	170.42395	1.00590
25	2.537	0.9944097	0.0055903	178.88065	1.00562
26	2.552	0.9946446	0.0053554	186.72887	1.00538
27	2.568	0.9948856	0.0051144	195.52725	1.00514
28	2.583	0.9951027	0.0048973	204.19429	1.00492
29	2.597	0.9952979	0.0047021	212.67133	1.00472
30	2.610	0.9954728	0.0045272	220.88943	1.00455
31	2.623	0.9956420	0.0043580	229.46338	1.00438
32	2.636	0.9958055	0.0041945	238.40755	1.00421
33	2.647	0.9959395	0.0040605	246.27467	1.00408
34	2.658	0.9960697	0.0039303	254.43153	1.00395
35	2.670	0.9962074	0.0037926	263.67236	1.00381
36	2.680	0.9963189	0.0036811	271.65601	1.00369
37	2.690	0.9964274	0.0035726	279.90479	1.00358
38	2.700	0.9965330	0.0034670	288.43164	1.00348
39	2.710	0.9966358	0.0033642	297.24683	1.00338
40	2.720	0.9967359	0.0032641	306.36035	1.00327
41	2.730	0.9968333	0.0031667	315.78247	1.00318
42	2.737	0.9968998	0.0031002	322.56421	1.00311
43	2.745	0.9969744	0.0030256	330.51367	1.00303
44	2.753	0.9970474	0.0029526	338.68042	1.00296
45	2.762	0.9971275	0.0028725	348.13257	1.00288
46	2.770	0.9971972	0.0028028	356.77979	1.00281
47	2.777	0.9972568	0.0027432	364.53979	1.00275
48	2.784	0.9973153	0.0026847	372.47925	1.00269
49	2.792	0.9973808	0.0026192	381.79492	1.00263
50	2.800	0.9974448	0.0025552	391.35986	1.00256

N	K10	P10	Q10	TP10	TQ10
51	2.808	0.9975075	0.0024925	401.19580	1.00250
52	2.814	0.9975535	0.0024465	408.75171	1.00245
53	2.822	0.9976137	0.0023863	419.06348	1.00239
54	2.828	0.9976580	0.0023420	426.97705	1.00235
55	2.834	0.9977015	0.0022985	435.07104	1.00230
56	2.841	0.9977514	0.0022486	444.71216	1.00225
57	2.847	0.9977933	0.0022067	453.16870	1.00221
58	2.853	0.9978345	0.0021655	461.78784	1.00217
59	2.860	0.9978818	0.0021182	472.09229	1.00212
60	2.867	0.9979280	0.0020720	482.63086	1.00208
61	2.872	0.9979606	0.0020394	490.33228	1.00204
62	2.878	0.9979990	0.0020010	499.73828	1.00200
63	2.883	0.9980304	0.0019696	507.72339	1.00197
64	2.889	0.9980676	0.0019324	517.49585	1.00194
65	2.893	0.9980921	0.0019079	524.12402	1.00191
66	2.898	0.9981222	0.0018778	532.54224	1.00188
67	2.903	0.9981520	0.0018480	541.11304	1.00185
68	2.908	0.9981812	0.0018188	549.82007	1.00182
69	2.913	0.9982101	0.0017899	558.68164	1.00179
70	2.918	0.9982386	0.0017614	567.71826	1.00176
71	2.923	0.9982666	0.0017334	576.89331	1.00174
72	2.928	0.9982942	0.0017058	586.24683	1.00171
73	2.932	0.9983160	0.0016840	593.84155	1.00169
74	2.937	0.9983429	0.0016571	603.47510	1.00166
75	2.941	0.9983642	0.0016358	611.32544	1.00164
76	2.946	0.9983904	0.0016096	621.26318	1.00161
77	2.950	0.9984111	0.0015889	629.37354	1.00159
78	2.954	0.9984316	0.0015684	637.57739	1.00157
79	2.958	0.9984518	0.0015482	645.89844	1.00155
80	2.963	0.9984767	0.0015233	656.46265	1.00152
81	2.967	0.9984964	0.0015036	665.04980	1.00151
82	2.970	0.9985110	0.0014890	671.57202	1.00149
83	2.974	0.9985302	0.0014698	680.36865	1.00147
84	2.979	0.9985540	0.0014460	691.55859	1.00145
85	2.982	0.9985681	0.0014319	698.38135	1.00143
86	2.986	0.9985867	0.0014133	707.57080	1.00142
87	2.990	0.9986051	0.0013949	716.88306	1.00140
88	2.994	0.9986233	0.0013767	726.34912	1.00138
89	2.998	0.9986412	0.0013588	735.93945	1.00136
90	3.002	0.9986589	0.0013411	745.65381	1.00134
91	3.005	0.9986721	0.0013279	753.05054	1.00133
92	3.008	0.9986851	0.0013149	760.52637	1.00132
93	3.012	0.9987023	0.0012977	770.58667	1.00130
94	3.015	0.9987150	0.0012850	778.23608	1.00129
95	3.019	0.9987319	0.0012681	788.58813	1.00127
96	3.023	0.9987485	0.0012515	799.06714	1.00125
97	3.027	0.9987650	0.0012350	809.71118	1.00124
98	3.030	0.9987772	0.0012228	817.80225	1.00122
99	3.033	0.9987893	0.0012107	825.93481	1.00121
100	3.037	0.9988052	0.0011948	836.97754	1.00120

TABLE 2 --Critical Values, Probablities and Return Periods for Two-tail Test for Detecting Outliers at 5 Percent Level

N	K5	P5	Q5	TP5	TQ5
4	1.689	0.9543902	0.0456098	21.92509	1.04779
5	1.869	0.9691886	0.0308114	32.45555	1.03179
6	1.996	0.9770331	0.0229669	43.54088	1.02351
7	2.093	0.9818255	0.0181745	55.02205	1.01851
8	2.172	0.9850723	0.0149277	66.98935	1.01515
9	2.237	0.9873569	0.0126431	79.09453	1.01280
10	2.294	0.9891048	0.0108952	91.78358	1.01101
11	2.343	0.9904353	0.0095647	104.55113	1.00966
12	2.387	0.9915068	0.0084932	117.74063	1.00857
13	2.426	0.9923669	0.0076331	131.00853	1.00769
14	2.461	0.9930725	0.0069275	144.35242	1.00698
15	2.493	0.9936665	0.0063335	157.89131	1.00637
16	2.523	0.9941821	0.0058179	171.88361	1.00585
17	2.551	0.9946293	0.0053707	186.19421	1.00540
18	2.577	0.9950169	0.0049831	200.67720	1.00501
19	2.600	0.9953388	0.0046612	214.53691	1.00468
20	2.623	0.9956420	0.0043580	229.46338	1.00438
21	2.644	0.9959034	0.0040966	244.10324	1.00411
22	2.664	0.9961391	0.0038609	259.00757	1.00388
23	2.683	0.9963517	0.0036483	274.10156	1.00366
24	2.701	0.9965434	0.0034566	289.30225	1.00347
25	2.717	0.9967061	0.0032939	303.59399	1.00330
26	2.734	0.9968715	0.0031285	319.63892	1.00314
27	2.751	0.9970293	0.0029707	336.62134	1.00298
28	2.768	0.9971799	0.0028201	354.59302	1.00283
29	2.781	0.9972904	0.0027096	369.05444	1.00272
30	2.794	0.9973969	0.0026031	384.16406	1.00261
31	2.808	0.9975075	0.0024925	401.19580	1.00250
32	2.819	0.9975913	0.0024087	415.15430	1.00241
33	2.833	0.9976943	0.0023057	433.71021	1.00231
34	2.846	0.9977863	0.0022137	451.74097	1.00222
35	2.858	0.9978684	0.0021316	469.12207	1.00214
36	2.869	0.9979411	0.0020589	485.69067	1.00206
37	2.880	0.9980116	0.0019884	502.91406	1.00199
38	2.890	0.9980738	0.0019262	519.14502	1.00193
39	2.900	0.9981341	0.0018659	535.94458	1.00187
40	2.910	0.9981928	0.0018072	553.33813	1.00181
41	2.919	0.9982442	0.0017558	569.53003	1.00176
42	2.925	0.9982777	0.0017223	580.60669	1.00173
43	2.937	0.9983429	0.0016571	603.47510	1.00166
44	2.945	0.9983852	0.0016148	619.26807	1.00162
45	2.954	0.9984316	0.0015684	637.57739	1.00157
46	2.960	0.9984618	0.0015382	650.10327	1.00154
47	2.970	0.9985110	0.0014890	671.57202	1.00149
48	2.978	0.9985493	0.0014507	689.31396	1.00145
49	2.985	0.9985821	0.0014179	705.25098	1.00142
50	2.993	0.9986187	0.0013813	723.96704	1.00138

N	K5	P5	Q5	TP5	TQ5
51	3.000	0.9986501	0.0013499	740.78125	1.00135
52	3.007	0.9986808	0.0013192	758.01807	1.00132
53	3.013	0.9987066	0.0012934	773.14355	1.00130
54	3.020	0.9987361	0.0012639	791.19141	1.00126
55	3.025	0.9987568	0.0012432	804.35400	1.00124
56	3.032	0.9987853	0.0012147	823.21948	1.00122
57	3.040	0.9988171	0.0011829	845.37012	1.00118
58	3.046	0.9988404	0.0011596	862.36011	1.00116
59	3.051	0.9988596	0.0011404	876.87305	1.00114
60	3.058	0.9988859	0.0011141	897.56128	1.00111
61	3.063	0.9989043	0.0010957	912.64819	1.00110
62	3.070	0.9989297	0.0010703	934.29932	1.00107
63	3.075	0.9989474	0.0010526	950.06592	1.00105
64	3.082	0.9989719	0.0010281	972.64844	1.00103
65	3.086	0.9989856	0.0010144	985.79321	1.00101
66	3.090	0.9989992	0.0010008	999.17896	1.00100
67	3.096	0.9990192	0.0009808	1019.58154	1.00098
68	3.101	0.9990356	0.0009644	1036.91064	1.00097
69	3.105	0.9990486	0.0009514	1051.07227	1.00095
70	3.110	0.9990645	0.0009355	1068.95288	1.00094
71	3.115	0.9990802	0.0009198	1087.24097	1.00092
72	3.121	0.9990988	0.0009012	1109.60400	1.00090
73	3.125	0.9991109	0.0008891	1124.77979	1.00089
74	3.130	0.9991260	0.0008740	1144.10889	1.00087
75	3.134	0.9991378	0.0008622	1159.76880	1.00086
76	3.138	0.9991494	0.0008506	1175.69824	1.00085
77	3.142	0.9991609	0.0008391	1191.81738	1.00084
78	3.148	0.9991780	0.0008220	1216.53345	1.00082
79	3.152	0.9991892	0.0008108	1233.34668	1.00081
80	3.157	0.9992030	0.0007970	1254.65259	1.00080
81	3.161	0.9992138	0.0007862	1271.96460	1.00079
82	3.164	0.9992219	0.0007781	1285.11792	1.00078
83	3.168	0.9992325	0.0007675	1302.98340	1.00077
84	3.172	0.9992430	0.0007570	1321.04053	1.00076
85	3.176	0.9992533	0.0007467	1339.28442	1.00075
86	3.180	0.9992636	0.0007364	1357.92920	1.00074
87	3.184	0.9992737	0.0007263	1376.76147	1.00073
88	3.188	0.9992837	0.0007163	1396.00708	1.00072
89	3.191	0.9992911	0.0007089	1410.56128	1.00071
90	3.194	0.9992984	0.0007016	1425.30078	1.00070
91	3.198	0.9993080	0.0006920	1445.19043	1.00069
92	3.202	0.9993176	0.0006824	1465.38696	1.00068
93	3.205	0.9993247	0.0006753	1480.77808	1.00068
94	3.208	0.9993317	0.0006683	1496.22900	1.00067
95	3.211	0.9993386	0.0006614	1511.86938	1.00066
96	3.214	0.9993455	0.0006545	1527.84033	1.00065
97	3.217	0.9993523	0.0006477	1543.86816	1.00065
98	3.220	0.9993590	0.0006410	1560.09058	1.00064
99	3.224	0.9993679	0.0006321	1582.00977	1.00063
100	3.228	0.9993767	0.0006233	1604.24707	1.00062


Percent Chance
X Units per 1000
(a)
(b)
(c)
(d)
$K_n = +3.09023231$
$K_n = -3.09023231$
$K_n = 0.0$
+2.50759
-2.50759
-3.037
10%
-3.228
5%
+3.037 for 10%
+3.228 for 5%
$\bar{X}$ =Arithmetic Mean
Sx=Standard Deviation
Cv=Coefficient of Variation
$(X/\bar{X}) = (1 \pm K_n Cv)$
$X(1000) = \bar{X} \pm K_n Sx$
FIGURE 1
Arithmetic Normal
Sample Size=100

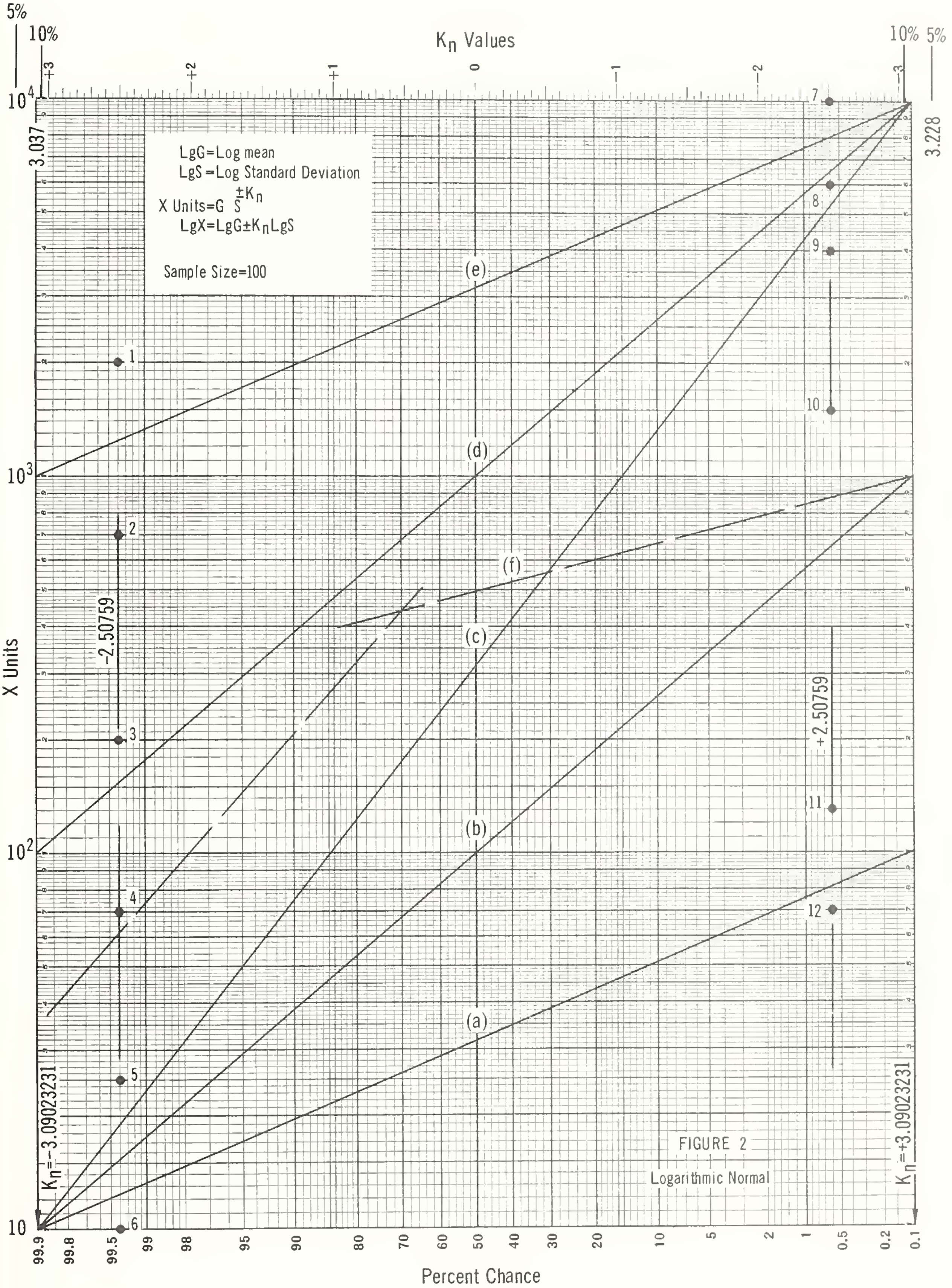


FIGURE 2

Logarithmic Normal

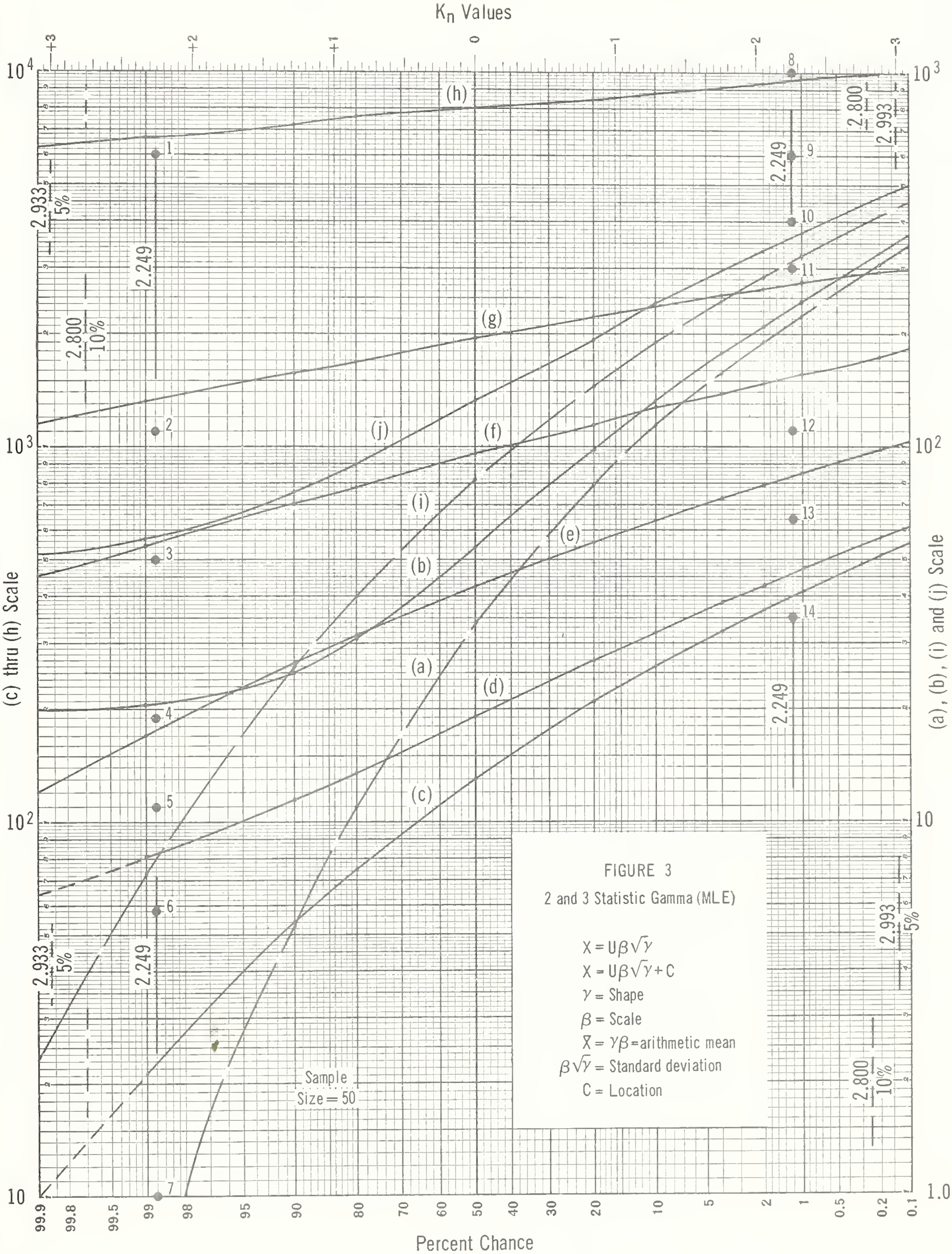


FIGURE 3

2 and 3 Statistic Gamma (MLE)

$X = U\beta\sqrt{\gamma}$

$X = U\beta\sqrt{\gamma} + C$

γ = Shape

β = Scale

$\bar{X} = \gamma\beta$ = arithmetic mean

$\beta\sqrt{\gamma}$ = Standard deviation

C = Location